UW-Madison Visit Day, 2018

Department of Biostatistics and Medical Informatics

Jacob M. Maronge

Department of Statistics
Department of Biostatistics and Medical Informatics

University of Wisconsin–Madison

https://jmmaronge.github.io/
jmmaronge
@jmmaronge

Introduction

About me

- 2nd year PhD student in statistics and biostatistics
- Working with Professor Paul Rathouz
- Research interests: clinical biostatistics, experimental design and longitudinal or clustered data, including biased sampling schemes

Goals

Why am I here?

I'd like to offer an example of some of the interesting work, **from a student prospective**, that you can be a part of here at UW.

Two Examples:

- Analysis of Speech Trajectories in Children with Cerebral Palsy (CP)
- Power Analysis for Longitudinal Data Using Decomposition

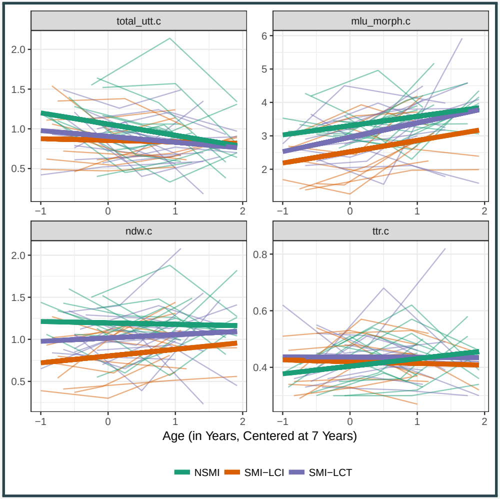
Analysis of Speech Trajectories

in children with CP

- Collaborators: Wisconsin Intelligibility, Speech, and Communication Laboratory (WISC Lab, PI: Professor Katie Hustad)
- CP is an umbrella term for a group of permanent movement disorders, which can also affect speech
- Clinicians categorize these children into 3 categories based off of speech abilities.

Question: Do data-driven methods regarding analysis of speech patterns collected on 35 children in the cohort suggested existence of these clinical categories?

Analysis of Speech Trajectories



Subject level data for 4 speech variables with group level fitted values from a multivariate longitudinal model

Power Analysis for Longitudinal Data

Using Decomposition

- Power analysis and sample size calculations are a crucial component of any study design.
- Many study designs lead to either complicated formulas or do not have closed form solutions
- We propose a method for calculating power in the situation of clustered or longitudinal data that leads to a solution that is both simple to calculate and gives valuable insight to how study design parameters affect power.

Setup

Model

Consider the situation where predictors are randomly observed. Our model of interest is,

$$Y_{ij} = \beta_0 + \beta X_{ij} + \epsilon_{ij}, \tag{1}$$

where $i = 1, \ldots, m$ denotes the study participant and $j = 1, \ldots, n$ denotes measurements within subject.

- $\mathrm{corr}(X_{ij}, X_{ij'}) = \tau_X > 0$
- $\mathrm{corr}(\epsilon_{ij}, \epsilon_{ij'}) = \tau_\epsilon > 0$
- $\mathrm{Var}(X_{ij}) = \mathrm{Var}(\epsilon_{ij}) = 1$

Decomposition

Results

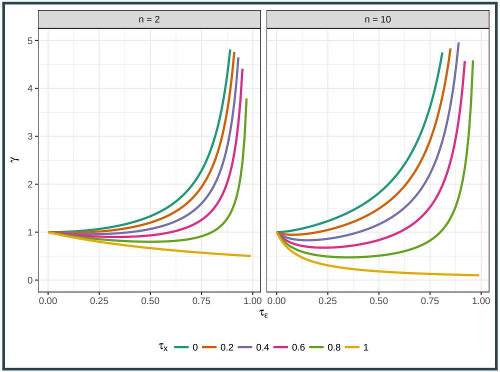
By decomposing the model into pieces which vary within-subject and between-subject, we are able to derive the correlation between X_{ij} and Y_{ij},

$$\rho^2_{\text{eff}} = \frac{\rho^2\gamma}{\rho^2\gamma + (1 - \rho^2)}. \tag{2}$$

Interesting pieces

- ρ denotes the correlation as if there were no correlation within subject.
- γ is a function of τ_X, τ_ϵ, and n (study design parameters).

Power Analysis for Longitudinal Data



- When $\tau_X = \tau_\epsilon$: equivalent to an independent random sample
- When $\tau_X = 1$: recover the classical result for longitudinal studies comparing treatment to control group

Conclusions

- I hope this shows good examples of exciting work **you** could do at UW
- Please reach out to me with any questions after your visit
- Special thanks to Professor Karl Broman for the template for these slides